

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MARYLAND**

**Equal Employment Opportunity
Commission**

v.

Performance Food Group, Inc.

*
*
*
*
*
*

Civil Action No. 13-1712

MEMORANDUM

Performance Food Group, Inc. (“PFG”) delivers food and food-related products through its foodservice distributors. One distributor, its Broadline Division,¹ manages around 20 distribution centers (“OpCos”) across the country. The Equal Employment Opportunity Commission (“EEOC”) alleges that PFG engaged in a pattern or practice of gender discrimination in the selection of operative positions at these distribution centers from 2004–2013. Now pending before the court is PFG’s motion to exclude the reports and testimony of the EEOC’s expert, Elvira Sisolak, and the EEOC’s second² motion to exclude the testimony and reports of PFG’s rebuttal expert, Stephen G. Bronars, Ph.D. The motions have been fully briefed and no oral argument is necessary. For the reasons stated below, the court will deny both motions.

FACTS

PFG employed “operatives,” defined by the EEOC as “[w]orkers who operate machine or processing equipment or perform other factory-type duties of intermediate skill level which can be mastered in a few weeks and require only limited training,”³ at its OpCos. The specific job titles at issue in this litigation are: (1) truck drivers; (2) selectors; (3) forklift operators; (4)

¹ Also known as “Performance Foodservice.”

² The EEOC’s first motion to exclude the report of Bronars (ECF 198) was denied without prejudice, in light of the additional testimony the EEOC could take from Bronars and the additional report to be filed by Sisolak. (ECF 220).

³ Job Patterns for Minorities and Women in Private Industry: A Glossary, <https://www.eeoc.gov/eeoc/statistics/employment/jobpat-eeo1/glossary.cfm> (last accessed March 12, 2020).

transportation or warehouse supervisors; and (5) an “other warehouse” category of miscellaneous nonselector warehouse jobs.

The EEOC’s expert, Elvira Sisolak, has submitted an expert report, a rebuttal report, and a supplemental rebuttal report. PFG’s rebuttal expert, Stephen G. Bronars, has submitted a rebuttal expert report and a supplemental rebuttal report. All reports concern Sisolak’s statistical tests showing a statistically significant gender disparity adverse to women in job offer rates in the five operative positions.

a. Sisolak’s Expert Report

Sisolak is a senior economist who works at the Office of General Counsel, EEOC. (ECF 256-1, Sisolak Report at 4). Her first report is dated August 7, 2017. Sisolak divided the applicant and hiring data provided by PFG into two time periods (2004–June 30, 2009, and July 1, 2009–2013), because the 2009–2013 period has more complete data. (*Id.* at 2–3). She also divided the “numerous Operative titles shown in the data into five job groups: Drivers, Forklift Operators, Selectors, Supervisors, and Other Warehouse positions.” (*Id.* at 2). PFG received 76,589 applications during the 2004–2009 period, (*id.* at 19), and 101,769 applications during the 2009–2013 period, (*id.* at 8).⁴ Between 2004–2009, PFG hired 9,863 applicants into operative jobs, (*id.* at 19), and between 2009–2013, PFG hired 5,051 applicants into operative jobs, (*id.* at 9). The largest group of applicants and hires was for the selector position, which is an entry-level position at the warehouse. (*Id.* at 8). The second largest group was for the driver position, which has some form of commercial licensing requirement and sometimes a preference for experience. (*Id.* at 9). The other job groups were much smaller, and the “other warehouse” category included daytime jobs similar to the selector position. (*Id.*).

⁴ This number does not include applications in which gender was not indicated, which were not included in the analyses. (Sisolak Report at 7). Additionally, the data for the 2004–2009 period, specifically 2006, may not be complete. (*Id.* at 19).

Sisolak compared the percentage of women applicants to the percentage of women selected. (*Id.* at 11 (2009–2013 period), 20–21 (2004–2009 period)). For each job group, she subtracted the actual female hires from the expected female hires, and calculated the “missed opportunities.” (*Id.* at 11 (2009–2013), 21 (2004–2009)). She also compared the selection rates for men and women for each of the job groups. (*Id.* at 12 (2009–2013), 21–22 (2004–2009)).

Sisolak then controlled for certain variables by conducting the analysis of selection rates by job group, location, and/or year. She did this by organizing the data into strata (e.g., one strata would consist of applicants for a selector position at a certain OpCo), and then aggregating the results to determine whether the selection rates were statistically significant. (*Id.* at 13–14 (2009–2013), 23 (2004–2009)). She used the Cochran-Mantel-Haenszel Procedure (“CMH”) to aggregate the statistical results; the test is “designed to test the null hypothesis that the true selection rates for women and men do not differ when the data is disaggregated by strata (level) and then recombined.”⁵ (*Id.* at 13). In total, Sisolak conducted CMH tests by job group; by job group and warehouse location; by job group, location, and year; and by job tracking number (“JTN”)⁶ (a unique requisition number for a specific job vacancy or vacancies). (*Id.* at 13–14 (2009–2013), 23 (2004–2009)). Each of these CMH tests showed a statistically significant gender disparity in selections. (*Id.* at 14 (2009–2013), 23 (2004–2009)).

⁵ Sisolak further explains the CMH test methodology in a declaration she provided. The CMH test aggregates the results of simpler tests, such as the chi square test or Fisher’s Exact test, which uses four values, in this case the number of male applicants, the number of female applicants, the number of male offers, and the number of female offers, to determine any disparity between the expected and actual offers for each gender. (ECF 261-8, Sisolak Dec. at 5–6). For example, to control based on differences in jobs, locations, requisition pools, points in time, and decisionmakers, the disparity can be calculated individually for each job tracking number (“JTN”). (*Id.*). To do this for each JTN and then determine if there is a correlation between offers and gender, the CMH test will generate the difference in actual and expected offers in each JTN and then combine these differences to determine if they are statistically significant. (*Id.*).

⁶ Job tracking numbers were not available for the 2004–2009 period, so this test was not conducted for that period. (Sisolak Report at 23 n.39).

Sisolak then controlled for Class A license (for driver applicants), experience, and whether the applicant completed an online application. Sisolak conducted a separate analysis of selection rates for driver applicants who did and did not have a Class A license. She applied a Fisher's Exact test⁷ to find that among the applicants who had a Class A license, there was a statistically significant disparity in selection rates adverse to women. (*Id.* at 14 (2009–20013), 23 (2004–2009)). To control for experience, Sisolak searched the applications to determine if the applicant had relevant prior experience, and sorted applicants based on having relevant experience for a job group, not having relevant experience, and missing information. (*Id.* at 15). For the 2009–2013 period, she compared the selection rates among applicants with experience by job group, and aggregated the results using CMH, finding a statistically significant gender disparity. (*Id.* at 16).⁸ Sisolak performed the same test among those applicants without experience, also finding a gender disparity not expected to occur by chance. (*Id.*). For the 2004–2009 period, there was not enough data on prior experience to control for it in the forklift and supervisor groups. (*Id.* at 24). Sisolak conducted “Fisher's Exact tests on the numbers of men and women hired and not hired as Selectors with relevant experience and without relevant experience and, separately, for Drivers with and without experience.” (*Id.*). Sisolak found statistically significant disparities among experienced and non-experienced applicants for the selector and driver positions in the 2004–2009 period. (*Id.*). For the 2009–2013 period (when PFG used online applications), some candidates completed a candidate profile on the PFG online application system, but not an online application.⁹ Sisolak used a Fisher's Test to compare the

⁷ The Fisher's Exact test is a measure of the discrepancy between the observed data and what would be expected by chance, and is designed to measure statistical significance in small sample sizes. *United States v. Hernandez-Estrada*, 749 F.3d 1154, 1164 (9th Cir. 2014).

⁸ It does not appear that Sisolak performed this analysis with respect to the “other warehouse” group.

⁹ Not every OpCo required applicants to complete the second part of the process, the online application. (Sisolak Report at 10).

selection rates of male and female applicants who did not complete an online application, finding a statistically significant hiring disparity adverse to female applicants. (Sisolak Report at 10).

b. Bronars' Expert Report

PFG provides the report of Stephen G. Bronars, dated December 18, 2017, to rebut Sisolak's report. Bronars does not provide a conclusion of his own regarding whether selection disparities based on gender occurred, but points out several flaws in Sisolak's report. Bronars states that Sisolak's conclusions are unreliable because she did not "model the hiring process at PFG." (*See* ECF 233-3, Bronars Report at 4). This is because, according to Bronars, Sisolak placed all applicants in one strata, regardless of whether they were actually competing against each other for a position, and ignores that each applicant did not have an equal chance of being selected, but might be more or less likely to be chosen based on their qualifications and who they were competing against. (*See, e.g., id.* at 9 & n.13 (falsely assumes equal chance), 16 ("artificially constructed candidate pools")). The CMH tests "artificially combine[] candidates and offers into composite annual pools" that do not account for whether those candidates were actually competing against each other for the same position or whether the candidates were eligible to receive offers for that position.¹⁰ (*Id.* at 22 (specifically regarding the CMH test by group/OpCo/year); *see id.* at 35 (2004–2009 data)). Further, some of the calculations and tests, including Sisolak's calculation of "missed opportunities" and her CMH test based on job group, included non-competitive requisitions, male-only requisitions, and requisitions for which no offer was made. (*Id.* at 11, 17).

¹⁰ For example, Bronars argues that to control for relevant prior experience with respect to the specific job opening candidates applied for, Sisolak should have "compare[d] the hiring outcomes of male and female candidates competing for the same position, while controlling for the quantity and quality of candidates' work experience." (*Id.* at 30).

An example of Bronars' criticism of the "artificial pools" is as follows: in her test controlling for online applications, Sisolak divided applicants based on whether they completed an online application. For some requisitions, however, only applicants who completed online applications were offered jobs. (*Id.* at 13). Therefore, for such a requisition ("requisition 1"), there would be no offers among the applicants who did not complete online applications. These "requisition 1" applicants would then be combined with other applicants from other requisitions who also did not complete online applications ("non-online application group"), and the gender disparity between group makeup of the non-online application group and offers would be calculated. Bronars' criticism, though, is that none of the "requisition 1" applicants would be eligible for any of the offers in the non-online application group, since the only requisition they applied to was offered to someone who completed an online application, and therefore this offer would instead be included in the "online application group." (*Id.*). According to Bronars, this way of analyzing the data "caused Ms. Sisolak to incorrectly measure expected job offer rates for the candidates she intended to study." (*Id.*). Because a slightly higher percentage of women failed to submit an online application than men, the failure to complete an online application could be a legitimate explanation for "missed opportunities" and, according to Bronars, was not adequately controlled for. (*Id.* at 14).

Bronars also criticizes Sisolak for aggregating her test results across all OpCos and job groups. He argues that such aggregation is inappropriate because of the difference in job requirements and hiring criteria across job groups and locations. (*Id.* at 15).

Bronars' other criticisms of Sisolak include that: she did not take into account the timing of applications in her CMH test based on job tracking number, (*id.* at 24); did not properly

categorize applicant's experience or use all available information,¹¹ (*id.* at 26–30 (2009–2013), 38–42 (2004–2009)¹²); and used job offers instead of job hires, when for some requisitions there were at least twice as many offers as hires, so Sisolak “overstated the number of hires that could have occurred and the magnitude of any shortfall in the hiring of women” and the alleged “missed opportunities” “includes offers to women that would have been rejected or rescinded,” (*id.* at 23).

c. Sisolak's Rebuttal Report

In her rebuttal report, dated March 6, 2018, Sisolak responds to Bronars' criticism of her aggregation of results, noting that such aggregation is necessary to prove a “pattern” of gender disparity. (ECF 256-9, Sisolak Rebuttal at 2). According to Sisolak, Bronars' methodological criticism is that Sisolak “should not have used CMH tests which took as their strata any population of candidates other than those candidates applying under the same job tracking number.” (*Id.* at 19). She argues that this position, that “one cannot aggregate data across any potential control variable,” is contrary to “routine analytical practice” and would be impossible for the 2004–2009 period, during which PFG did not use job tracking numbers and it is not clear which applicant applied to which specific position. (*Id.* at 20). Sisolak explains why she did not use the Career Builder resumes, and the issues with matching applicants to the PDF applications.

¹¹ This could be important because it appears that men were more likely to have relevant prior experience than women. (*Id.* at 30).

¹² For example, for the 2004–2009 applicants, the applications provided by PFG did not include past experience. The EEOC hired a vendor to key in experience from PDF application files that were separately provided by PFG to the EEOC, but Bronars argues that his sample of some applications show that some applicants were erroneously matched to the PDF application, or could have been matched to their PDF application but were not. (*Id.* at 38–40). Additionally, PFG provided the EEOC with a CareerBuilder file with resumes. Bronars randomly sampled 100 records from the CareerBuilder resume file and was able to match 37 resumes to applicants and categorize the experience of these applicants. (*Id.* at 40–41 & n.61). He surmises that Sisolak could have accounted for the experience for a significant portion of the 2004–2009 applicants if she matched the applications with the CareerBuilder resumes. (*Id.* n.61).

(*Id.* at 11). She also notes that Bronars' criticism of her report is speculative, as he does not know if the outcome would have changed, since he did not perform tests of his own. (*Id.* at 3).

Sisolak also conducted additional tests to respond to Bronars. First, she conducted a CMH test that controlled for online applications, prior work experience, and Class A licenses (for driver applicants) at the same time, finding the results also statistically significant. (*Id.* at 3, 6–7). Second, she conducted CMH tests using hires instead of offers, and found that the results were essentially the same. (*Id.* at 25).

d. Bronars' Supplemental Report

In his supplemental report, dated April 16, 2018, Bronars conducted additional tests “to assess how Ms. Sisolak’s tests of gender disparities in job offer rates would have changed had she reported test results separately by job group and” OpCo. (ECF 233-5, Bronars Rebuttal Report at 1). He first conducted separate CMH tests “for each of 81 combinations of job groups and OpCos that had at least one ‘competitive’ requisition [male and female candidates and at least one offer] in 2009–2013 using job requisitions as strata.” (*Id.* at 3). He found no significant disparity in job offer rates adverse to females at any OpCo for the positions of forklift operator, warehouse supervisor, and other warehouse position; and found no significant disparities in offers for drivers in 15 of 22 OpCos, and no significant disparities in offers for selectors in 8 of 23 OpCos. (*Id.* at 3–4).

Second, Bronars “used a multiple regression methodology and Ms. Sisolak’s data to estimate job offer models for each job group at each OpCo while accounting for job requisition and her simple controls for prior experience.” (*Id.* at 6). Bronars found no significant differences in job offer rates for the driver, forklift operator, warehouse supervisor, or other

warehouse positions at any OpCo. (*Id.* at 6–7). He found no significant disparities in offer rates for female selector applicants in 9 out of the 23 OpCos. (*Id.* at 7).

e. Sisolak’s Supplemental Rebuttal Report

In response to Bronars’ rebuttal, Sisolak submitted a supplemental rebuttal, dated September 28, 2018. She argues that Bronars divided the data into unreasonably small groups so that results could never be statistically significant. (ECF 233-6, Sisolak Supplemental Rebuttal at 5–7). For example, Bronars calculated no statistically significant disparities in 44 location/job groups that had female applicants but zero female hires. (*Id.* at 7). Sisolak argues that Bronars relied on an inappropriate form of regression, Ordinary Least Squares (“OLS”), that is not appropriate when there are only two values (hired/not hired), which is shown by the fact that his analysis produces some probabilities of getting the job as less than zero. (*Id.* at 16). Even using the OLS regression test that Bronars used, however, relying on the job tracking numbers to take location into account and controlling for experience, Sisolak found that gender disparities in selection rates were statistically significant in all jobs combined, and for the driver, selector, and other warehouse job groups. (*Id.* at 17). Based on the OLS regression analysis, selection rates for women applying to the forklift and supervisor positions were lower than for men, but the difference was not statistically significant. (*Id.* at 17). Sisolak also reached the same results using another regression test, logistic regression. (*Id.* at 16–17).

STANDARD OF REVIEW

Rule 702 of the Federal Rules of Evidence, which governs the admissibility of expert testimony, states:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if: (a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue; (b) the testimony is based

on sufficient facts or data; (c) the testimony is the product of reliable principles and methods; and (d) the expert has reliably applied the principles and methods to the facts of the case.

The party seeking to introduce expert testimony has the burden of establishing its admissibility by a preponderance of the evidence. *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579, 592 n.10 (1993). A district court is afforded “great deference . . . to admit or exclude expert testimony under *Daubert*.” *TFWS, Inc. v. Schaefer*, 325 F.3d 234, 240 (4th Cir. 2003) (citations and internal quotation marks omitted); *see also Daubert*, 509 U.S. at 594 (“The inquiry envisioned by Rule 702 is . . . a flexible one . . .”). “In applying *Daubert*, a court evaluates the methodology or reasoning that the proffered scientific or technical expert uses to reach his conclusion; the court does not evaluate the conclusion itself,” *Schaefer*, 325 F.3d at 240, although “conclusions and methodology are not entirely distinct from one another.” *General Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997). In essence, the court acts as gatekeeper, only admitting expert testimony where the underlying methodology satisfies a two-pronged test for (1) reliability and (2) relevance. *See Daubert*, 509 U.S. at 589. To be admissible, however, “the expert testimony need not be irrefutable or certainly correct.” *Young v. Swiney*, 23 F. Supp. 3d 596, 611 (D. Md. 2014) (internal citation and quotation omitted). “In other words, the Supreme Court did not intend the gatekeeper role to supplant the adversary system or the role of the jury: [v]igorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.” *Id.* (internal citations and quotations omitted).

ANALYSIS

I. Motion to exclude Sisolak

Bronars’ rebuttal of Sisolak’s expert report mostly involves criticisms of the placement of

applicants into “artificial” statistical pools and the aggregation of selection rates across positions, OpCos, and years. Additionally, some criticisms relate to the failure of Sisolak to control for certain factors (or to control properly for those factors). Bronars also criticizes the inclusion in some tests of noncompetitive requisitions or requisitions in which only male candidates were included.¹³

Bronars’ criticism, even if valid, does not require exclusion of Sisolak’s expert testimony and reports. First, the use of CMH tests has been accepted in other cases,¹⁴ and Sisolak also applied Bronars’ preferred method, multiple regression analysis, in her supplemental rebuttal report. Although Bronars argues that Sisolak failed to model the hiring process by constructing artificial pools, he stated in his deposition that he was not sure whether there is a difference in statistical analysis between artificially constructed and organically given pools. (ECF 233-4, Deposition of Bronars at 216:4–12). Whether CMH or regression analysis is more appropriate for this data is a question of fact for the jury, since both Sisolak and Bronars have different

¹³ PFG also argues that in Sisolak’s rebuttal test, controlling for job tracking number, online application, and experience, she overestimated the likelihood that women would receive an offer. This is because her test divided the applicants into statistical pools or strata based on whether they completed an online application and whether they had experience. At least one of the pools or strata, however, had 10 male applicants and 0 female applicants and no job offers, so it was dropped from the CMH analysis. By dropping unsuccessful male applicants, “Sisolak’s approach distorts the chances of receiving an offer for both male and female applicants.” (ECF 283, PFG Reply in Supp. of Mot. in Limine at 11). It is not clear the extent to which this happened (e.g., in how many instances there were male-only strata with no offers that were dropped from the analysis). It appears this criticism applies only to Sisolak’s rebuttal test controlling for job tracking numbers and experience/online applications. Additionally, it does not appear that this criticism is explicitly made in Bronars’ rebuttal report.

¹⁴ See *Bazile v. City of Houston*, 858 F. Supp. 2d 718 (S.D. Tex. 2012) (two experts agreed that Mantel-Haenszel test was the most “methodologically appropriate analysis” to aggregate historical data as it “allows statisticians to investigate the consistency of data trends over time while avoiding errors due to aggregation,” *id.* at 741, and finding a prima facie case of disparate impact, in part, because the “Mantel–Haenszel analysis confirms that the historical patterns of 4/5 Rule [the EEOC’s rule for determining disparate impact] violations are statistically significant,” *id.* at 766); *Oliver v. VSoft Corp.*, No. 1:09-CV-0185-CAP-WEJ, 2010 WL 11505776, at *10 n.25 (N.D. Ga. Feb. 2, 2010), report and recommendation adopted, No. 1:09-CV-185-CAP-WEJ, 2010 WL 11506873 (N.D. Ga. Feb. 23, 2010) (“In its own research, the Court identified four federal cases recognizing the Mantel-Haenszel method in the context of discrimination claims” and collecting cases, although noting elsewhere that “the statistical significance attached to the aggregation method is disputed,” *id.* at *11) (unreported cases are cited for the soundness of their reasoning, not for any precedential value). The cases refer to “Mantel-Haenszel” rather than “Cochran-Mantel-Haenszel.” It appears that these tests are the same, based on the EEOC’s references to *Bazile* and *Oliver*, and PFG does not state anything to the contrary in its reply.

opinions that appear to be well-reasoned and reliant on scholarly sources. Second, cases have found that aggregation across job positions, locations, and time is appropriate in pattern or practice cases, given what the plaintiff must prove.¹⁵ See *EEOC v. Texas Roadhouse, Inc.*, 215 F. Supp. 3d 140, 157 (D. Mass. 2016) (“Given the nature of [pattern or practice] claims, aggregating data for a nationwide view is not improper or unduly prejudicial.”). Bronars’ own tests in his supplemental expert report show that it may be problematic not to aggregate, as certain subsets of positions and locations may not have enough statistical power to generate meaningful results.

Sisolak attempted to control for factors such as experience, whether the applicant completed an online application, and whether the applicant had a Class A license (for drivers) in subsequent tests, to show that they did not account for the findings of gender disparity. To the extent PFG argues that Sisolak did not properly control for experience, such as by not including additional information from sites like CareerBuilder, it may address this on cross-examination. See *Texas Roadhouse, Inc.*, 215 F Supp. 3d at 155 (“failing to use a perfect set of variables that incorporates all relevant factors or excludes all potentially irrelevant variables is not a means for rejecting an expert’s analysis.”). Importantly, it appears that many of Bronars’ criticisms, for example failing to properly measure experience or control for the timing of the application, could affect men and women equally, and may not alter the findings of gender disparity. Therefore, even if Bronars’ criticisms are valid, it is not clear how they reflect on the ultimate question.

¹⁵ PFG cites to *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 356–57 (2011), for the proposition that “[w]ithout a uniform policy or uniform decision-maker(s) across OpCos, Sisolak’s analysis of aggregated data has no foundation.” (ECF 256, Mot. in Limine to Exclude Sisolak at 32). *Dukes*, however, involved a question of class certification, and not a motion to exclude. Additionally, whether there was an unofficial uniform policy at PFG of not hiring women for operative positions is a question of fact. Finally, *Dukes* was a proposed class action for all female Wal-Mart employees. *Dukes*, 564 U.S. at 346. Here, the pattern or practice allegations concern only operative positions in the Broadline division of PFG, positions that are likely more similar to each other than the positions at issue in *Dukes*.

Similarly, PFG can question Sisolak as to her treatment of non-competitive and male-only requisitions. In particular, there appears to be a reasonable dispute as to how to treat non-competitive requisitions, which tended to be male only. That PFG appears to have created certain requisitions to hire a preselected (and usually) male candidate may be data that should be considered in the statistical analysis. (See Sisolak Supplemental Rebuttal at 11).

Therefore, the court finds Sisolak's reports and testimony to be relevant and reliable, and will deny the motion to exclude them.

II. Motion to exclude Bronars

The EEOC argues that Bronars' criticisms are not relevant or reliable because he merely speculates that certain factors might affect the outcome of Sisolak's tests, without showing that they would. Specifically, for many of the criticisms, he does not state how they would affect the ultimate finding of gender disparity. As to his second report, the EEOC argues that Bronars' findings are misleading, as many of his statistical pools did not have sufficient statistical power. The EEOC also argues that Bronars misstates the opinion of the EEOC's expert in another case, Dr. David Neumark, regarding the benefits of multiple regression over CMH.

Although "[t]he court's function is more limited when evaluating rebuttal expert testimony offered by the defendant," it must still "determine threshold admissibility." *Samuel v. Ford Motor Co.*, 112 F. Supp. 2d 460, 469 (D. Md. 2000). Therefore, "a rebuttal expert is still subject to the scrutiny of *Daubert* and must offer both relevant and reliable opinions." *Funderburk v. S.C. Elec. & Gas Co.*, 395 F. Supp. 3d 695, 716 (D.S.C. 2019).

Here, Bronars' opinions meets the standards for relevance and reliability. First, Bronars' rebuttal would be helpful to the factfinder because it demonstrates some limitations of Sisolak's methodology: for example, his testimony might show that multiple regression analysis, rather

than CMH, is more appropriate in these circumstances. His testimony also explains problems with aggregating across position and location, as the data might show a statistically significant selection disparity as to all operatives in the Broadline division when, in reality, such disparity only existed in a few positions or locations.¹⁶

Second, Bronars has sufficiently explained his criticisms so that they are reliable. The cases cited by the EEOC are distinguishable. In *Rembrandt Social Media, LP v. Facebook, Inc.*, No. 1:13-CV-158, 2013 WL 12097624, at *2 (E.D. Va. Dec. 3, 2013), the court excluded rebuttal testimony when the expert gave “no sources for this opinion, and provide[d] no reason to believe her opinion is based on a reasoned explanation.” Here, Bronars explains why he believes Sisolak failed to consider certain relevant factors, inappropriately aggregated the data, and inappropriately used CMH. *Id.* His criticisms of Sisolak’s data are testable, as evidenced by the fact that Sisolak performed additional tests to respond to some of Bronars’ criticisms.

The EEOC also points to *In re Ethicon, Inc.*, MDL No. 2327, 2016 WL 4493501, at *3 (S.D. W.Va. Aug. 25, 2016), where the court excluded rebuttal testimony when it was unclear what material the expert reviewed to reach her opinions, and reviewed only five samples of tissue explants without explaining her methodology in selecting and reviewing the samples. Although Bronars did rely on samples of applications to form his opinion that Sisolak mismatched some applicants with their applications or did not obtain all available experience data, Bronars used a larger number of random samples and explained how he reached his

¹⁶ It is true that some of Bronars’ testimony is speculative, as he does not know if considering certain factors would affect the finding of gender disparity. For example, the failure to consider the timing of applications (e.g. if they came in after the last offer was made) would seemingly only affect the gender disparity results if women were more likely than men to submit their applications after a position was filled, which PFG has not shown. Still, given that the EEOC has the burden to prove a pattern or practice of discrimination and this may impact their ability to meet that burden, the court finds the testimony relevant. Some courts have held that a rebuttal expert may identify flaws in another expert’s report even without showing that the flaws affected the results, and that doing so is not unduly speculative. *See Aviva Sports, Inc. v. Fingerhut Direct Mktg., Inc.*, 829 F. Supp. 2d 802, 834–35 (D. Minn. 2011) (collecting cases).

opinion. (ECF 233-3, Bronars Report, Appendix C and Appendix D).¹⁷ This is also unlike *Eghnayem v. Boston Scientific Corp.*, 57 F. Supp. 3d 658, 673–74 (S.D. W.Va. 2014), which the EEOC also cites to, in which the court found an expert’s opinions based on a sample of pathology reports, chosen by the plaintiffs, to be unreliable, since the plaintiffs did not explain how they chose the reports. Here, Bronars, rather than the parties, chose the samples, and he states he did so randomly.¹⁸

In Bronars’ second report, he analyzed 81 separate statistical pools, based on job position and OpCo, and found no significant disparities with respect to certain positions and certain locations. The EEOC argues that Bronars’ calculations are misleading because many of the pools did not have sufficient statistical power (so that even with zero female hires the disparity was not statistically significant), and because Bronars gives equal weight to the pools even though some had hundreds more applicants than the others.

Here, Bronars’ methodology is clearly laid out in his report, so it is testable (and in fact, Sisolak did test his methodology, including by calculating the statistical power of his data). (Sisolak Supplemental Rebuttal at 9). To the extent the EEOC contends that Bronars’ findings are misleading, the EEOC may question him about that issue on cross-examination.¹⁹

¹⁷ The EEOC also cites to *Crawford v. Newport News Industrial Corp.*, No. 4:14-cv-130, 2017 WL 3222547, at *6–7 (E.D. Va. July 28, 2017). In *Crawford*, however, the expert relied on adjustments to the data by a management official, and the expert did not know how the official made these adjustments, but believed that some of the adjustments were based on his memory. *Id.* at *7. Here, Bronars relied on the applicant data provided by PFG, and there is no indication the data was altered or based on personal memory.

¹⁸ The EEOC also cites to *Faust v. Comcast Cable Commc'ns Mgmt., LLC*, No. CIV.A. WMN-10-2336, 2014 WL 3534008, at *4 (D. Md. July 15, 2014), which excluded an expert report for multiple reasons, including because his summary was based on a “remarkably small” sample size of plaintiffs, and the majority of data he included was outside the relevant time period of the case. In contrast, Bronars’ random sample of resumes is not meant to be a summary of the data, but is relevant to show that some of the resumes may have been mischaracterized and that Sisolak could have included additional data in her analysis.

¹⁹ The EEOC also argues that Bronars’ application of his criticisms in his supplemental report to the 2004–2009 period is not reliable because Bronars did not perform his statistical tests for the 2004–2009 period. Bronars states in his report that “[m]y criticisms of Ms. Sisolak’s methodology apply to her studies of the 2004–2009 period, as well.” (Bronars Supplemental Report at 2). Because he does not attempt to extrapolate the results based on his tests

Additionally, Bronars' supplemental report is relevant in that it relates to the gender disparities among job positions and OpCos, and it demonstrates that results of the statistical tests may vary depending on the portion of the data analyzed and the type of statistical analysis performed.

Finally, with respect to David Neumark's testimony, Bronars and Sisolak disagree as to the meaning of his statements regarding CMH in his testimony as an expert in a separate case. Sisolak and Bronars may dispute the general acceptance of CMH and multiple regression analysis in the scientific community, since that is relevant to whether CMH and multiple regression are reliable tests. The court notes, though, that Neumark's opinions on the use of CMH and multiple regression analysis in an unrelated case are likely of little (if any) relevance to this case, and any reference to Neumark is relevant only in the context of general scientific acceptance.

Therefore, the court finds Bronars' testimony and reports to be relevant and reliable, and will deny the motion to exclude them.

CONCLUSION

For the reasons stated above, the court will deny the motions to exclude as to Elvira Sisolak and Stephen G. Bronars. A separate order follows.

3/18/20
Date

/s/
Catherine C. Blake
United States District Judge

for 2009–2013 to 2004–2009, but only seeks to criticize the methodology, much like he did in his first report, the court finds it sufficiently reliable and relevant to be admissible.